

2017（第十二届）城市发展与管理大会，中国·海口  
分论坛：数据驱动的可持续城市

# 浅谈中国的政府数据和非政府数据

Governmental Data and Non-governmental Data in China

毛其智

清华大学，2017年7月27日

# 清华大学成立中国经济社会数据中心

(China Data Center, Tsinghua University, CDC)

2009年11月22日，清华大学成立了中国经济社会数据中心。长期以来，中国政府部门、学术界及其他单位积累了大量数据，但都很少向公众开放。中心的任务就是整理、利用国家统计局及相关机构的数据，进行政策研究；通过与大数据企业合作，挖掘和分析大数据，为政府统计提供新的数据来源，为国家高层决策提供咨询服务等。中心的一手数据都将免费向研究者开放。

2016年7月21日，清华大学校长邱勇与国家统计局局长宁吉喆签署协议，合作共建“清华大学中国经济社会数据研究中心”。原国家统计局副局长、清华大学兼职教授许宪春出任中心主任，中心设立学术指导委员会和执行委员会。



# 什么是政府数据？

**数据(数值):** data, 数据就是数值, 即人们通过观察、实验或计算得出的结果。数据有很多种类, 一般包括数字数据和模拟数据。(数据可以是静态的, 或动态的, 并不断有新型数据出现)。

**政府统计数据:** 统计工作活动过程中所取得的反映国民经济和社会现象的数字资料以及与之相联系的其他资料的总称, 包括国家统计局调查、部门统计调查和地方统计调查(分为: 公开数据、内部数据、涉密数据)。

**政府数据**是国家执政能力的宣示, 以国家、部门和地方统计机构所公布的数据为准。**政府数据应真实反映中国的经济社会发展状况。**

# 中国统计“四大工程”

## ▶ 马建堂局长提出四大工程构想

2010年12月,在全国统计工作会议上,马建堂局长提出了四大工程的概念。

今后几年,是加快统计工作科学发展的关键时期,是深化统计改革、推进统计建设的攻坚阶段。抓住统计事业发展的重要战略机遇期,奋力提高统计能力、统计数据质量和政府统计公信力,全面推进统计工作规范统一、改革创新、公开透明,充分发挥统计在国家宏观调控和经济社会发展管理中的基础性作用,非常重要的一个抓手就是要加快建设基本单位名录库、企业一套表制度、数据采集处理软件系统和联网直报系统等互相联系、共为整体的四大工程。



对统计调查制度进行统一电子化设计和布置,具备数据统一管理、录入、审核、编辑、报送、汇总等功能,性能优良、便于操作的数据采集处理软件系统。

## ▶ 四大工程的主要任务

### 第一大工程

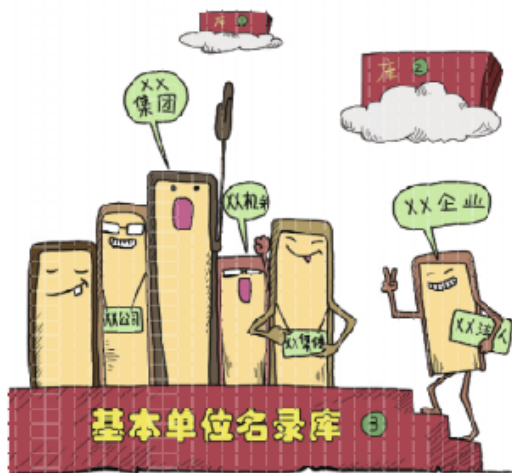
The First Program

#### 基本单位名录库

Business Register

建设真实完整、及时更新的基本单位名录库。

坚持“全国统一管理、专业分工协作、地方分级负责、各方共同参与、信息资料共享”的原则,按照“统一标准、一库在线、分级维护、及时更新”的模式,遵循“各项统计调查必须使用统一的名录库作为调查单位库或抽样框,不在名录库中的单位不得列入专业统计调查范围”的要求,以经济普查资料为基础,充分利用部门行政记录和专业统计信息,健全基本单位统计标准,完善基本单位名录库维护更新规范,分步建设“三上”企业调查单位库和“三下”企业及非企业单位名录库,最终建成一个全国统一完整、不重不漏、真实准确、及时更新的基本单位名录库。



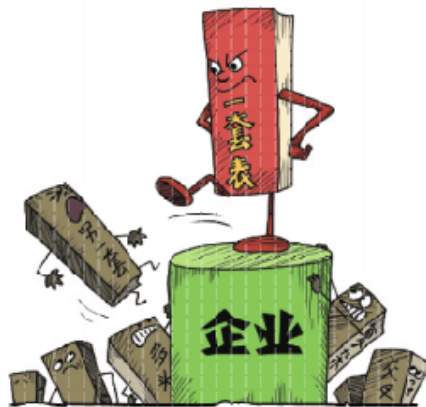
### 第二大工程

The Second Program

#### 企业(单位)一套表制度

Integrated Questionnaire for Enterprises

建立统一规范、方便填报的企业(单位)一套表制度。按照“统一设计、统一标准、统一调查单位、统一布置”的原则,将对企业分散实施的各项调查整合统一到一起,统一布置报表,统一采集原生性指标数据,统一不同专业报表中相同指标的涵义、计算方法、分类标准和统计编码,建立既能有效满足各级党委政府、各类经济体和社会公众统计需求,又能满足专业统计和国民经济核算需要,便于企业填报、减轻企业和基层统计机构负担的统一规范的企业一套表制度。



### 第三大工程

The Third Program

#### 数据采集处理软件系统

Data Collection and Processing Software

建设功能完善、统一兼容的数据采集处理软件系统。按照“功能完善、方便使用、标准统一、友好兼容”的总体要求,以规范的统计业务流程为依托,以解决现有数据采集处理软件多乱为重点,以满足各项统计调查数据采集处理、实现不同专业数据共享为目标,建设能够



### 第四大工程

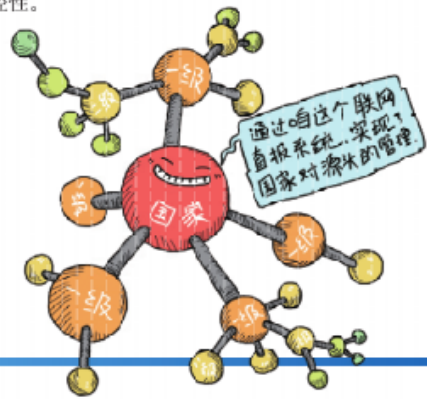
The Fourth Program

#### 联网直报系统

Online Reporting System

建设安全畅通、便捷高效的联网直报系统。

在信息化硬件设施、数据采集处理软件系统和原始数据库建设的基础上,尽快实现调查对象和调查人员通过互联网直接向全国数据管理中心报送原始数据、各级统计机构在线共享的工作模式,转变基层统计队伍工作重点,从过去繁重的数据收集汇总、报表填报转向对原始数据的核查和企业基础统计工作的督导,有效消除可能存在的中间环节对统计数据的干扰,提高数据汇总效率和生产过程的透明度与可控性。



# 什么是非政府数据？

数据是国家政治经济与社会文化生活领域重要的基础性资源，是进行各种科学研究的基础。当今世界已进入信息化、互联网和大数据的时代，一个重大改变就是信息的扁平化，**人人都是数据的生产者和使用者**，对数据的采集、加工、汇总、综合、分析、预测，已成为全社会各方面的共同需求，并比任何时候都要迫切。

**非政府数据**就是指除国家、政府部门和地方统计机构所公布的数据之外，来自海内外各行各业以及公民个体的数据。**非政府数据的采集、提供和使用应遵守国家的相关法律法规。**

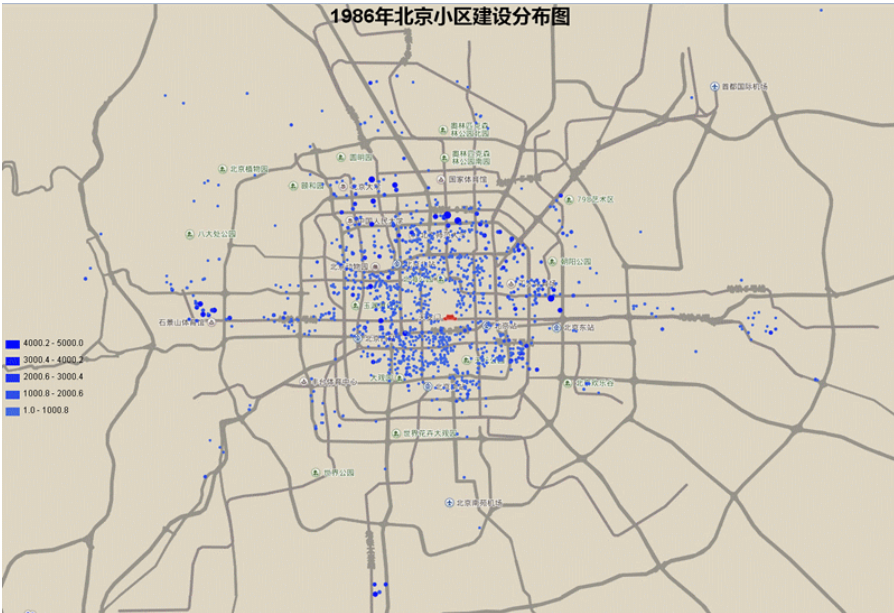




**数量规模：100+社区高手**

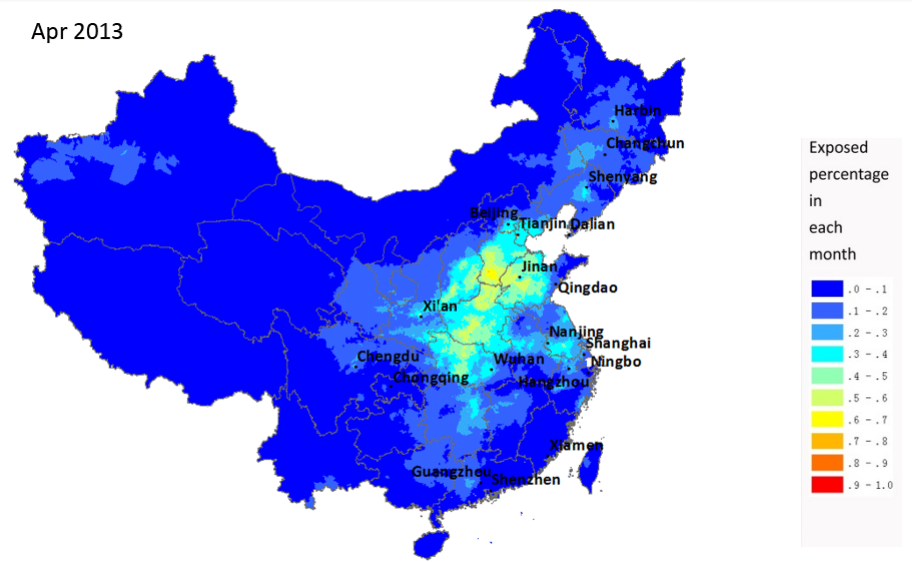
**分布于全国40+城市**





1986-2015年北京市区居住小区建设增长示意图

Apr 2013



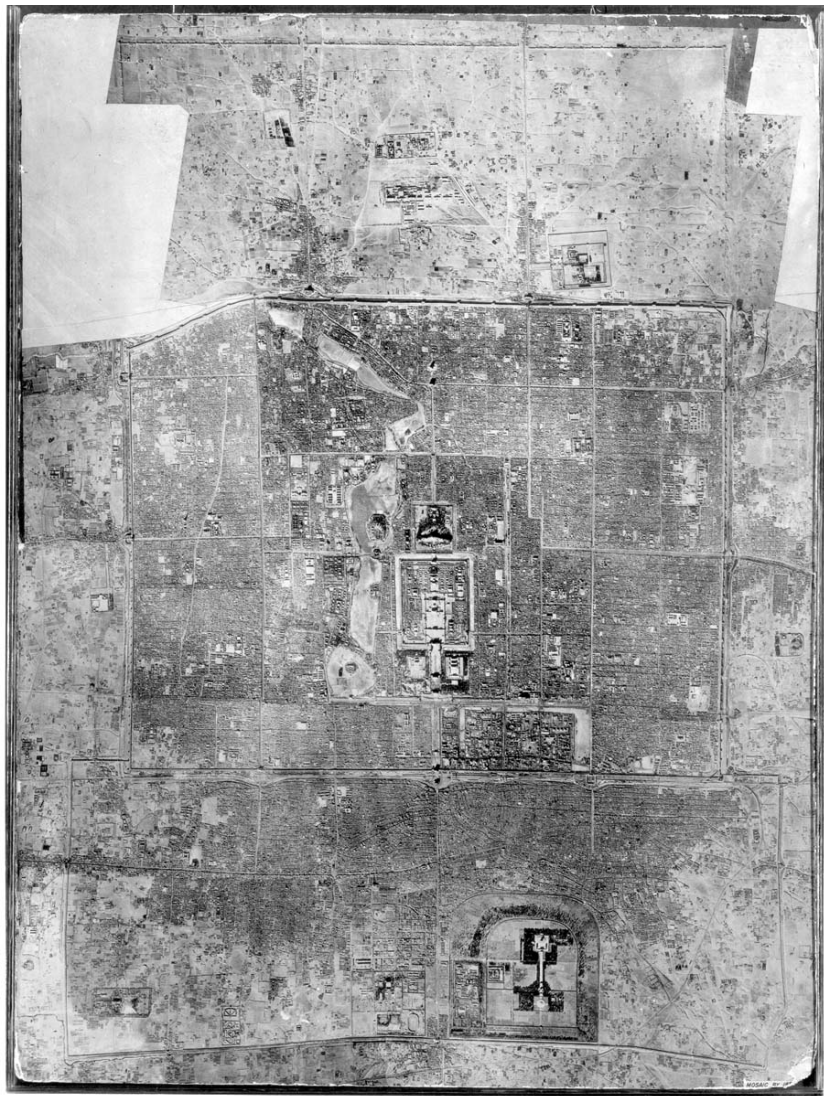
摩拜在广州的一天



6-10点通州向北京市区早高峰通勤动态



以前：数据短缺（eg. 北京城市航拍，从1943年美军侦察到8301工程）





## 今天：数据“爆炸” (eg. MB, GB, TB, PB, EB, ZB, ...)

2017年5月26日，由国家发改委、工信部，国家互联网信息办、贵州省人民政府共同主办的以“数字经济引领新增长”为主题的2017中国国际大数据产业博览会于在贵阳开幕。中国科学院院长白春礼在开幕式上说：

目前，全球数据总量每年都以倍增的速度增长，预计到2020年将达到**44万亿GB**，中国数据量到2020年将占全球数据总量的近20%。

据估计，2015年全球大数据核心产业产值超过300亿美元，**潜在规模超过8000亿美元**。

目前中国科学院已存储科学大数据规模达到**数百个PB**。

**以城市数据为例，获取之目的：**

城乡规划、建设、管理， ...

城市研究，衣食住行，生老病死，安居乐业， ...

城市经营，房地产市场， ... ..

**城市数据规模庞大、内容繁多**

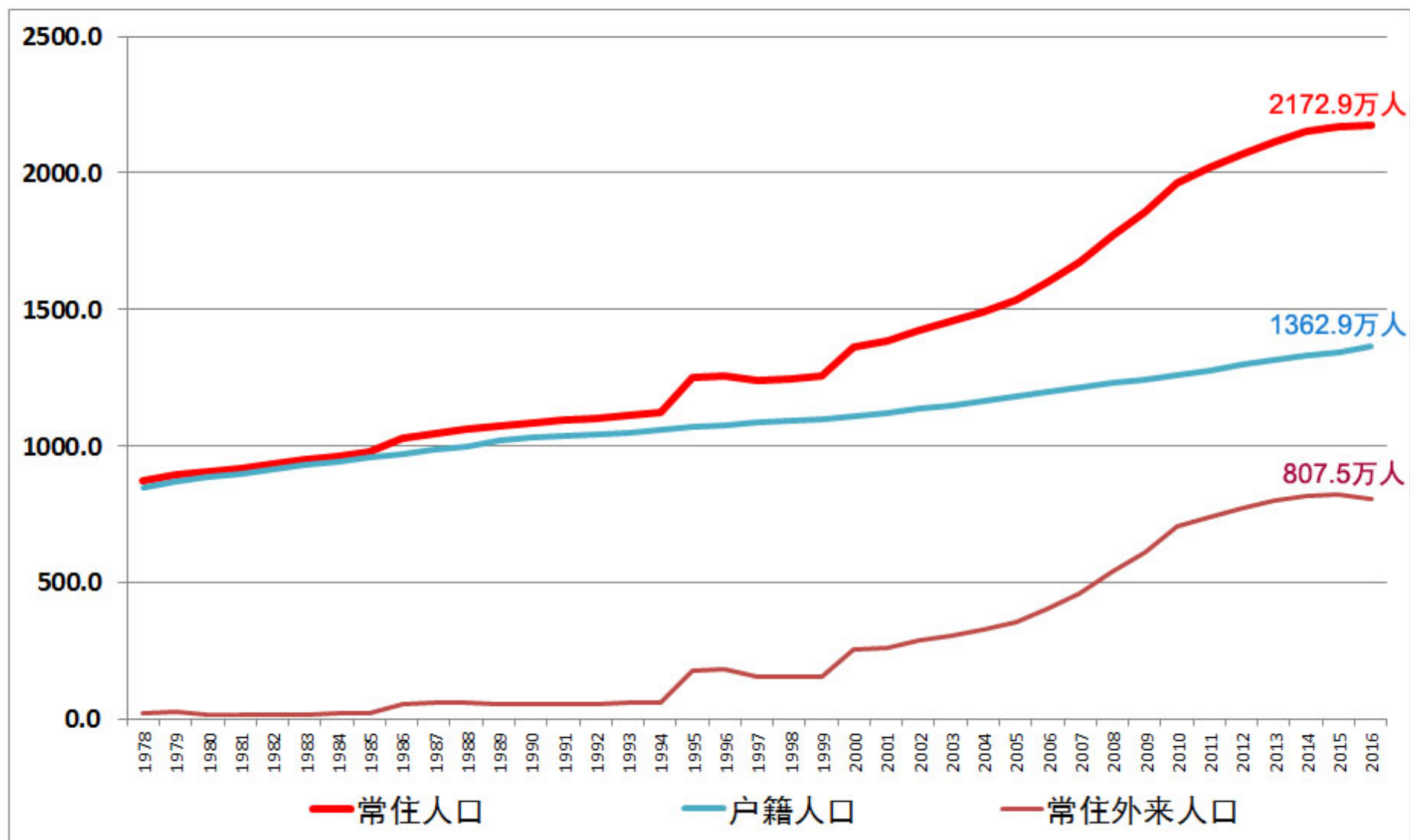
**其中与规划相关的数据只是“沧海一粟”**

**对数据的学习、认识和理解**

**应与我们的时代同步成长**



## 北京市人口发展趋势 (1978-2016年)



# 第一次全国地理国情普查公报 (2017年4月)

地理国情现状 1:16 000 000

标准时点：2015年6月30日

图1

## 地理国情要素分类

- 种植土地
- 林草覆盖
- 水域
- 荒漠与裸露地
- 铁路与道路
- 居民地与设施

地理国情要素分类



基础地理要素



注1：第一次全国地理国情普查公报  
注2：本公报数据截至2015年6月30日  
审图号：GS(2016)1632号

## 地理国情是空间规划体系最基础的数据之一





## 有关数据的几个问题

- 数据来源：**国家统计局，地方统计局，其它政府部门，社会数据，商业数据，来自海外的数据， ...
- 数据定义：**人口，建成区，建筑面积，收入，雾霾， ...
- 采集时间：**同时（遥感），基准时（普查），时段， ...
- 数据精度：**采集精度，标识精度，需求精度， ...
- 数据误差：**采集误差，合理误差，其它误差， ...
- 数据真伪：**真实数据，保密数据，加工数据， ...
- 数据解释：**基础解释（采集与观察），应用解释（原理）

....

# 有关信息和数据的法律及政府文件（举例）

《**中华人民共和国统计法**》，1983年12月8日六届全国人大常委会第三次会议通过，2009年6月27日修订，自2010年1月1日起施行。

《**中华人民共和国统计法实施细则**》，1987年1月19日国务院批准，1987年2月15日国家统计局发布。根据2005年12月16日国务院的决定修订，2006年2月1日起施行。

《**中华人民共和国标准化法**》，1988年12月29日七届全国人大常委会第五次会议通过，自1989年4月1日起施行。2017年2月22日国务院常务会议审议通过《**中华人民共和国标准化法（修订草案）**》，决定将草案提请全国人大常委会审议。

《**中华人民共和国著作权法**》，1990年9月7日七届全国人大常委会第十五次会议通过，2010年2月26日通过第二次修正，自2010年4月1日起施行。

《**企业信息公示暂行条例**》，2014年7月23日国务院第57次常务会议通过，自2014年10月1日起施行。

《**国务院关于促进信息消费扩大内需的若干意见**》，2013年8月8日，提出我国信息消费市场规模庞大，但信息消费面临基础设施支撑能力有待提升、产品和服务创新能力弱、市场准入门槛高、配套政策不健全、行业壁垒严重、体制机制不适应等问题。加快促进信息消费，是一项既利当前又利长远、既稳增长又调结构的重要举措。

《**最高人民法院、最高人民检察院关于办理侵犯公民个人信息刑事案件适用法律若干问题的解释**》，2017年3月20日最高法第1712次会议、2017年4月26日最高检第63次会议通过，自2017年6月1日起施行。

《**中华人民共和国测绘法**》，1992年12月28日七届全国人大常委会第二十九次会议通过，2017年4月27日修订，自2017年7月1日起施行。



## 小结

- 1 政府数据是中国今天各类数据的重要基础**
- 2 非政府数据是政府数据社会化的根本保障**
- 3 政府数据与非政府数据相互促进共同发展**

# 谢谢各位聆听!

联系地址:

毛其智

北京·清华大学建筑学院

邮编: 100084

<http://www.arch.tsinghua.edu.cn/chs/>

Email: [qizhi@tsinghua.edu.cn](mailto:qizhi@tsinghua.edu.cn)

Tel.: 010-62783328, 62782622